


DATA NOTE

Open Access



Draft genome sequence of *Paenibacillus* sp. EZ-K15 isolated from wastewater systems

Waleed S. Mohammed^{1,2}, Elvira E. Ziganshina¹, Elena I. Shagimardanova³, Natalia E. Gogoleva³ and Ayrat M. Ziganshin^{1*} 

Abstract

Objectives: *Paenibacillus* species, belonging to the family *Paenibacillaceae*, are able to survive for long periods under adverse environmental conditions. Several *Paenibacillus* species produce antimicrobial compounds and are capable of biodegradation of various contaminants; therefore, more investigations at the genomic level are necessary to improve our understanding of their ecology, genetics, as well as potential biotechnological applications.

Data description: In the present study, we describe the draft genome sequence of *Paenibacillus* sp. EZ-K15 that was isolated from nitrocellulose-contaminated wastewater samples. The genome comprises 7,258,662 bp, with a G+C content of 48.6%. This whole genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession PDHM00000000. Data demonstrated here can be used by other researchers working or studying in the field of whole genome analysis and application of *Paenibacillus* species in biotechnological processes.

Keywords: *Paenibacillus* sp., Genome sequencing, Draft genome assembly and annotation, Wastewater

Objective

Paenibacillus species, belonging to the family *Paenibacillaceae*, are rod-shaped Gram-positive or Gram-variable endospore forming aerobic or facultatively anaerobic bacteria, which are able to survive for long periods under adverse environmental conditions. Bacteria belonging to the genus *Paenibacillus* can be isolated from a wide range of environments including humans, animals, plants and the environment [1, 2]. Many species of *Paenibacillus* genus synthesize antimicrobial compounds that can be used as pesticides as well as in medicine, and many species produce enzymes important in bioremediation related technologies. *Paenibacillus* strains can be successfully applied for contaminants removal from a variety of wastewater systems. Also, several *Paenibacillus* strains can be involved in hydrolysis of cellulose and hemicellulose, lignin depolymerization and degradation of various textile dyes, polyvinyl alcohol, diesel fuel, bitumen,

polycyclic aromatic hydrocarbons, benzene and other compounds [2]. Hence, more studies at the genomic level are important to clarify our understanding of their ecology, genetics, as well as potential biotechnological applications. Thus, *Paenibacillus* sp. EZ-K15 was isolated from nitrocellulose-contaminated wastewater systems, Kazan, Republic of Tatarstan, Russia [3]. These industrial wastes produce high levels of wastewaters polluted with multifarious dissolved chemical compounds and nitrocellulose particles. Therefore, isolation of bacteria which are able to transform various adverse pollutants and their genome analysis are of high importance for the creation of effective bioremediation strategies [4, 5].

Data description

Paenibacillus sp. EZ-K15 was isolated from nitrocellulose-contaminated wastewater environments, Kazan, Republic of Tatarstan, Russia [3]. The bacterium optimally grown on Luria agar at + 30 °C had been cultivated for 1–2 days. Genomic DNA of the bacterial strain EZ-K15 was extracted using a FastDNA spin kit (#116540600; MP Biomedicals) and a Super FastPrep-1 homogenizer (#116011500; MP Biomedicals) as detailed

*Correspondence: a.ziganshin06@fulbrightmail.org

¹ Department of Microbiology, Institute of Fundamental Medicine and Biology, Kazan (Volga Region) Federal University, Kremlyovskaya str. 18, Kazan 420008, Russia

Full list of author information is available at the end of the article

in the manufacturer's protocol. Concentration and purity (A260/A280) of the obtained bacterial genomic DNA were measured using a NanoDrop 2000 spectrophotometer (#ND-2000; Thermo Fisher Scientific) and then stored at -20°C . The bacterial strain *Paenibacillus* sp. EZ-K15 was morphologically identified and then confirmed by PCR amplification applying universal primers UniBac27f, Bakt_805R and Univ1492r, followed by sequencing using an ABI PRISM 3130xl Genetic Analyzer (#4359571; Applied Biosystems) and phylogenetic analysis (16S rRNA gene sequence, 1437 bp, BLAST identity of 99% to *Paenibacillus lautus*; Table 1). In order to perform whole genome analysis, DNA was fragmented using a Q800R2 Sonicator (#Q800R2-110; Qsonica), and DNA library was then prepared using a NEBNext Ultra DNA Library Prep Kit for Illumina (#E7370S; New England Biolabs) according to the manufacturers' protocols. Both efficiency of DNA fragmentation and DNA library preparation were monitored using a 2100 Bioanalyzer (#G2939BA; Agilent) and a High Sensitivity DNA kit (#5067-4626; Agilent). Further sequencing was performed using a high-throughput Illumina MiSeq platform (#SY-410-1003; Illumina) at Joint KFU-Riken Laboratory, Kazan Federal University (Kazan, Russia) by a MiSeq Reagent Kit v2PE 500 cycles (#MS-102-2003; Illumina). Sequence reads quality was assessed using PRINSEQ lite version 0.20.4 [6], the genome was assembled using Velvet version 1.2.10 [7], and the ordering of contigs was achieved using Mauve version 2.4.0 [8]. The whole genome sequence of *Paenibacillus* sp. EZ-K15 was annotated using the Rapid Annotation System Technology server [9]. The rRNA and tRNA genes were identified using RNAmmer 1.2 [10] and tRNA scan-SE 1.23 [11], respectively. Comparison of the genomic feature of *Paenibacillus* sp. EZ-K15 with several *Paenibacillus* strains was performed using data obtained from EzBioCloud database [12].

The draft genome sequence of *Paenibacillus* sp. EZ-K15 composed of 36 contigs ranging from 512 to 911,265 bp with a total size of 7,258,662 bp, a G+C content of 48.6% and N50 of 242,001 bp. The Rapid Annotation System Technology server predicted 6682 coding sequences where 2551 coding sequences (39%) were annotated as seed subsystem features and 4131 coding sequences (61%) were annotated as outside of the seed subsystem. In total 4560 and 2122 coding sequences were assigned as non-hypothetical and hypothetical, accordingly. The

genome was shown to encode at least 3 rRNAs and 66 tRNAs. The strain *Paenibacillus* sp. EZ-K15 possesses a substantial number of genes responsible for denitrification and nitrate/nitrite ammonification (e.g., for nitrate released during nitrocellulose denitration) as well as for metabolism of aromatic compounds, including genes involved in benzoate, gentisate and some other compounds biodegradation. Numerous genes responsible for resistance to toxic compounds, including arsenic, mercury, cadmium, as well as chromium compounds, were additionally detected. This resistant strain may have future usefulness in bioremediation of various sites.

Limitations

Current data is based on the draft level genome sequence, due to which exact length of the genome, synteny, number of rRNA and repetitive elements cannot be reported. In addition, whether the genome consists of any plasmid/s or extra-chromosomal DNA cannot be certainly predicted.

Authors' contributions

WSM and EEZ conducted experiments, performed genome analysis, interpretation of the data and drafted the manuscript. EIS and NEG carried out Illumina sequencing and revised the manuscript. AMZ supervised the project, designed the study, performed genome analysis and professionally revised the manuscript. All authors read and approved the final manuscript.

Author details

¹ Department of Microbiology, Institute of Fundamental Medicine and Biology, Kazan (Volga Region) Federal University, Kremlyovskaya str. 18, Kazan 420008, Russia. ² Department of Biotechnology, Faculty of Agriculture, Al-Azhar University, Cairo 11651, Egypt. ³ Laboratory of Extreme Biology, Institute of Fundamental Medicine and Biology, Kazan (Volga Region) Federal University, Kazan 420021, Russia.

Acknowledgements

We thank Dr. Rushan Agzamov for support during the sampling.

Competing interests

The authors declare that they have no competing interests.

Availability of data materials

The data described in this Data note can be freely and openly accessed at DDBJ/ENA/GenBank. Please see Table 1 for details and links to the data.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Table 1 Overview of data files/data sets

| Label | Name of data file/data set | File types (file extension) | Data repository and identifier (DOI or accession number) | License |
|-------------|------------------------------|-----------------------------|---|---------|
| Data file 1 | Whole genome shotgun project | FASTA | DDBJ/ENA/GenBank (accession PDHM00000000) | CC-BY |
| Data file 2 | 16S rRNA gene sequence | FASTA | GenBank (accession MG050738) | CC-BY |

Funding

The reported study was funded by the Russian Foundation for Basic Research [Grant No. 16-34-60093 mol_a_dk].

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 20 October 2017 Accepted: 5 December 2017

Published online: 12 December 2017

References

- Sáez-Nieto JA, Medina-Pascual MJ, Carrasco G, Garrido N, Fernandez-Torres MA, Villalón P, Valdezate S. *Paenibacillus* spp. isolated from human and environmental samples in Spain: detection of 11 new species. *New Microbes New Infect.* 2017;19:19–27.
- Grady EN, MacDonald J, Liu L, Richman A, Yuan ZC. Current knowledge and perspectives of *Paenibacillus*: a review. *Microb Cell Fact.* 2016;15:203.
- Ziganshina EE, Ibragimov EM, Ilinskaya ON, Ziganshin AM. Bacterial communities inhabiting toxic industrial wastewater generated during nitrocellulose production. *Biologia.* 2016;71:70–8.
- Khilyas IV, Ziganshin AM, Pannier AJ, Gerlach R. Effect of ferrihydrite on 2,4,6-trinitrotoluene biotransformation by an aerobic yeast. *Biodegradation.* 2013;24:631–44.
- Ziganshin AM, Ziganshina EE, Byrne J, Gerlach R, Struve E, Biktagirov T, Rodionov A, Kappler A. Fe(III) mineral reduction followed by partial dissolution and reactive oxygen species generation during 2,4,6-trinitrotoluene transformation by the aerobic yeast *Yarrowia lipolytica*. *AMB Express.* 2015;5:8.
- Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics.* 2011;27:863–4.
- Zerbino DR. Using the Velvet *de novo* assembler for short-read sequencing technologies. *Curr Protoc Bioinform.* 2010;11(11):5.
- Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. Reordering contigs of draft genomes using the Mauve aligner. *Bioinformatics.* 2009;25:2071–3.
- Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST server: rapid annotations using subsystems technology. *BMC Genomics.* 2008;9:75.
- Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 2007;35:3100–8.
- Lowe TM, Eddy SR. tRNA scan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 1997;25:955–64.
- Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, Chun J. Introducing EzBioCloud: a taxonomically united database of 16S rRNA and whole genome assemblies. *Int J Syst Evol Microbiol.* 2017;67:1613–7.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

